Modeling and Optimizing High Pressure Liquid Chromatography (HPLC) Columns for the Separation of Biopharmaceuticals

BEE/MAE 4530
Computer Aided Engineering—Applications to Biomedical Processes

Group 4

Alex Huang
Dantong Huang
Jie Yuan

Table of Contents

Executive Summary

One of the most critical steps in the production of pharmaceuticals is the separation of the desired compound from reaction byproducts and environmental contaminants. Among the most sensitive of these methods is High Pressure Liquid Chromatography (HPLC), in which an initial mixture of compounds is forced by high pressure fluid flow through a column packed with a porous solid medium. Size and charge interactions with the solid phase cause the compounds to elute at different times from the column.

The performance of an HPLC column is highly dependent on properties such as the length, ambient temperature, inlet pressure, and solid medium porosity. The ideal parameters are conventionally determined by purchasing and physically testing a series of columns, which can be prohibitive in cost, time, and materials. Thus there currently exists a pressing need for computer models to simulate the separation of two or more compounds in order to expedite the onerous process of physical optimization.

This study sought to simulate the physical phenomena that underlie the elution process in an HPLC column, and optimize the conditions such that species separation and purity are maximized. The computing software COMSOL was used to model the involved physics, which comprised the flow of a mobile phase through a porous matrix, modeled by the Navier-Stokes Brinkman equation; the diffusion and dispersion of two solutes in the matrix, modeled by the general mass transfer equation; and the effect of external heating on the materials' behavior, modeled by the general heat equation. The geometry of the HPLC column consisted of an axisymmetric two-dimensional tube filled with a uniformly distributed porous matrix.

This model column was evaluated by simulating the separation of creatine and creatinine, two closely-related molecules involved in muscle tissue energetics. Once the model was tailored to a high degree of accuracy in comparison with experimental data, the column and species parameters were optimized. The optimal geometry for the separation of creatine and creatinine by HPLC, was a column of diameter 1.05 mm and length 78.4 mm, with a packed bed of spherical particles 5 μm in diameter. The optimal column temperature for this particular situation was found to be lower, at 15℃, as this slightly increases peak resolution but also elution time.

Though concentration plots derived from this model corroborated experimental elution absorbance plots with relatively high fidelity, lingering issues remain, including the unexpectedly small influence of temperature on elution characteristics. Future models may seek to correct this calculation error by including a less steep concentration gradient at the inlet at initial time points. Additionally, variations in column heating were found to have a very small effect on the diffusion of the solute bands, so the external temperature was excluded from the optimization process. The successful implementation of this model indicates that HPLC chromatography can be feasibly represented by computer modeling, and more specific models can reduce the time and material costs of extensive physical testing.

Introduction

Background

Pharmaceutical companies require on average 12 years and 350 million dollars to develop a drug from the laboratory to the pharmacy shelf (New Drug Approval Process, 2013). Since numerous guidelines regarding drug safety are closely enforced by the Food and Drug Administration, pharmaceutical companies often spend vast resources on identifying suitable testing methods during the production process. A typical method employed by companies is High Pressure Liquid Chromatography (HPLC). Liquid chromatography refers to the separation of a mixture of compounds. These compounds are dissolved in a fluid mobile phase, which carries them through a solid stationary phase that will cause the compounds to travel at different speeds, resulting in the separation of different species. Because of its unique abilities to separate, identify, and quantitate a variety of compounds, HPLC is one of the most powerful tools in the pharmaceutical industry (Waters, 2013).

Figure 1 shows a typical HPLC setup. A pump is used to generate the specific flow rate as well as the composition of mobile phases. The injector in the autosampler is responsible for sending the sample of interest into the HPLC system. Different compounds in the sample have different affinities with this stationary phase inside, resulting in different traveling velocities within the porous matrix of the column. As these compounds elute from the outlet of the column, their concentrations are measured using ultraviolet spectroscopy. Compounds detected in the fluid are represented by peaks in a chromatogram, a depiction of the elution profile of each compound over time. Depending on when a compound elutes from the column, the purified form can be collected downstream of the detector for further analysis or sent to waste if it is no longer needed after the study.

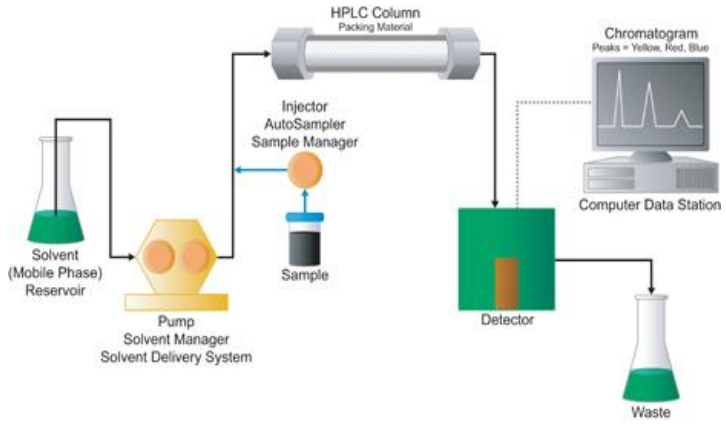

Figure 1. Diagram of a HPLC system. Samples mixed into a solvent are pumped at high pressure into the porous matrix of the HPLC column, through which substances travel at different speeds depending on their size and electrochemical properties. Purified samples are identified by absorbance measurements. Adapted from Waters (2008).

An ideal separation of two species by an HPLC column is characterized by a large difference in elution time between the species. Also, each peak should be relatively narrow; otherwise, a large amount of fluid must be collected to isolate the majority of the material, diluting the sample itself. Column dimensions such as diameter, length, packing material and packing particle size can have profound effects on the resolution of different compounds. To control the elution time, the user will also need to choose an appropriate temperature for the column chamber and an appropriate flow rate. The best HPLC method may require balancing a tradeoff between several opposing factors, such as achieving a low elution time as well as a high peak separation. The balance of these attributes for a given reaction is very specific to the given materials and situation, so testing is always required before a column is run on a large scale. Physically testing a series of columns to determine the right properties for a particular application is almost always labor intensive, time consuming, and expensive.

Problem statement

The goal of this study is to mitigate the HPLC laboratory testing process by developing a HPLC column computer model to simulate column behavior. Though laboratory testing is irreplaceable, a successful computer model can provide the user with a first pass solution, indicating whether or not certain combinations of parameters will work before the user invests in the actual experimentation. However, to be considered as a reasonable alternative to conventional testing, the computer model must simulate the physical processes to a high degree of accuracy.

In order to narrow down the various HPLC applications in analytical chemistry, reverse phase HPLC was considered in this project. The separation of the biomolecules creatine and creatinine is often conducted using this method. These two molecules are very similar in structure, and both are involved in ATP generation in tissues (MacNeil, 2005). Creatine is synthesized naturally in the kidneys and liver and is transported to the muscles. In times of low energy requirements, creatine undergoes a reversible reaction to phosphocreatine. Phosphocreatine builds up in the muscle tissue until a rapid surge of energy is needed, at which point phosphocreatine is converted back to creatine and ATP is generated. Both phosphocreatine and creatine degrade into creatinine, which is removed at a steady rate by the kidneys.

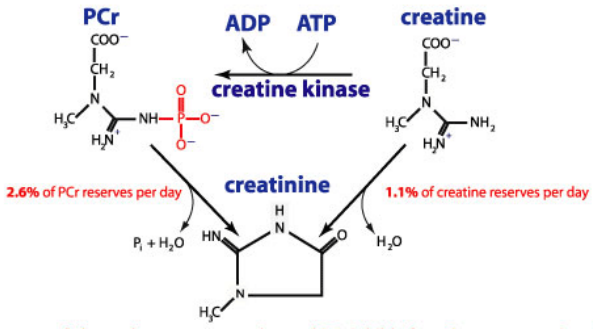

Figure 2. The creatine–phosphocreatine-creatinine reaction for ATP generation in muscle cells. The separation of creatine and creatinine is an often-performed HPLC process in medical research and pharmaceutical industry (Adapted from NSN, 2013).

The isolation and measurement of creatine and creatinine are very relevant in medicine. Since creatine is a popular supplement, doctors are interested in measuring the rate of creatine excretion from blood to identify possible adverse health effects from over consumption. Conversely, concentration measurements of creatinine in urine are often used to evaluate kidney function. Because of these applications, the separation of creatine and creatinine is a high demand process, and HPLC is a well-established method for performing it. For this reason, the properties of the separation process have been detailed extensively and it is an ideal subject for the HPLC model in this study. This separation process will serve as the standard for validation of the mathematical model developed, as well as the parameters for subsequent optimization analysis.

Design Objectives

The goal of this study is to model the separation of creatine and creatinine through a HPLC column using COMSOL Multiphysics. Specific objectives include:

1. By mimicking the column dimensions and molecular properties presented in a column evaluation report (Faulkner et al., 2013) of a porous graphite column, a chromatography model was sought which could very accurately reproduce the species elution profiles in the experimental chromatogram. This model would then be used for further analysis.
2. Following the modeling process, various column parameters were varied to achieve the optimal separation of creatine and creatinine. Separation quality was measured by the time at which the maximum concentration of each peak occurs, and by the width of each peak at half of the maximum concentration. The parameters to be investigated include:
    - Flow rate through the column;
    - Column length;
    - Column diameter;
    - Column heating temperature.

Optimal values discovered by this method were compared to the values listed by Faulkner et al., to determine how closely the physical model corresponds to the ideal model obtained using COMSOL.

Problem Schematic

The column was modeled as a 2D axisymmetric rectangle to simulate a cylinder as seen in Figure 3. A constant velocity was specified at the top inlet for the flow of solution at 20℃. Atmospheric pressure was specified at the outlet. At the axis of symmetry, all fluxes were assumed to be zero, and on the right boundary, a constant temperature was specified at 30℃. The no slip condition was also assumed at the surface of the outer boundary. The column had a radius of 1.05 mm and an axial length of 100 mm. Regarding initial conditions, it is assumed that there are no solutes in the column at t = 0s; but at the initial inflow, a constant concentration of 0.572 mol/m^3 of both solutes flows into the column for 1 second.

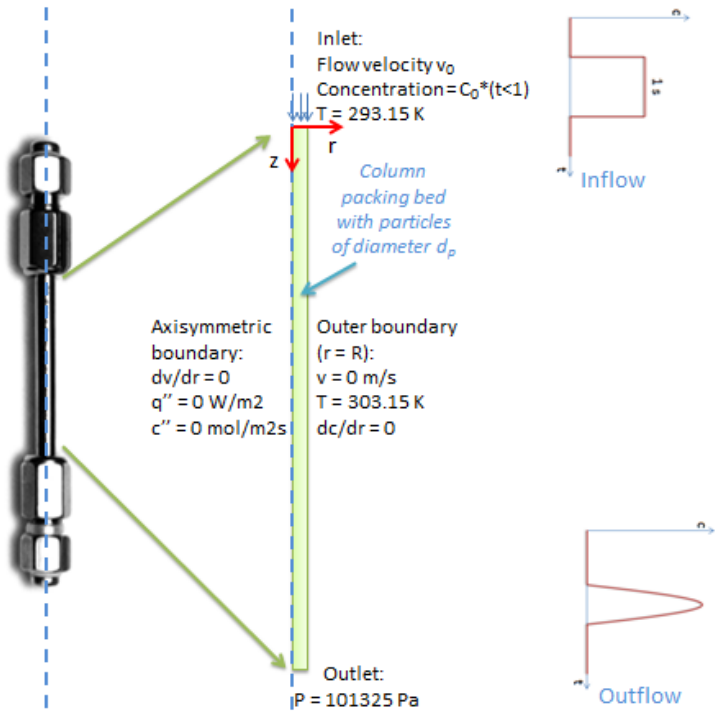

Figure 3. Column dimensions and boundary conditions of the column in the proposed HPLC model. A typical HPLC column is shown on the leftmost (adapted from Cobert Associate, 2013). The modeled column has a diameter of 2.1mm and a length is 100mm, with a particle diameter of 5μm. Two plots on the right show the expected concentration of each species at the inlet and outlet.

Results and Discussion

Full Solution

The goal of this project is to simulate a HPLC elution process, which is used to separate creatine and creatinine on an advertised column from the manufacturer's website—Hypercarb Reverse Phase HPLC Column by Thermo Fisher Scientific (Faulkner, 2013). In this project, the elution profile of a compound is determined by evaluating the average concentration at the end of the column as a function of time. The mobile phase flowing through the column is assumed to have properties similar to those of water and at it is at steady state.

Since flow is driven by high pressure to overcome the high compaction inside the column, a significant pressure drop is expected along the column axis, shown in Figure 4A. At a constant flow rate of 2 mL/min as specified by the column manufacturer, a pressure drop of more than 4×10^5 Pa is required. A flow rate of 2 mL/min corresponds to a linear velocity of 9.624×10^{-4} m/s, as shown in Figure 4B. Because the column contains a porous matrix under high pressure, the velocity profile is not expected to be a normal parabola, as in an empty pipe. Instead, velocity is a constant value at the interior, but drops sharply to zero at the wall due to the no slip boundary condition.

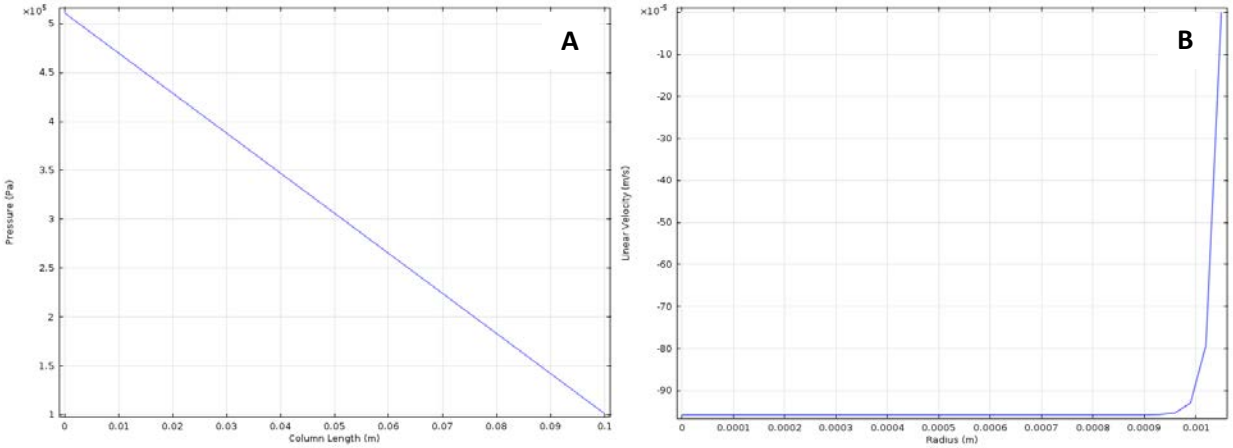

Figure 4. Steady state pressure [A] and velocity [B] profiles. [A] At a constant flow rate of 2 mL/min, a pressure drop of 4×10^5 Pa is required for the driven flow. [B] The linear velocity is 9.624×10^{-4} m/s, and no slip is assumed at the wall of the column (negative values are shown because flow is downward).

The HPLC column is constantly being heated at 30℃ at the wall during the experiment. The steady state 2D temperature profile of the column is shown in Figure 5. At the inlet of the column, temperature of the porous media is assumed to equilibrate with the temperature of the flowing water (20℃ or 293K). As water flows through the column, it only cools down the very beginning (~1.5mm) of the porous matrix, which is kept at a constant temperature (30℃) throughout the rest of the length.

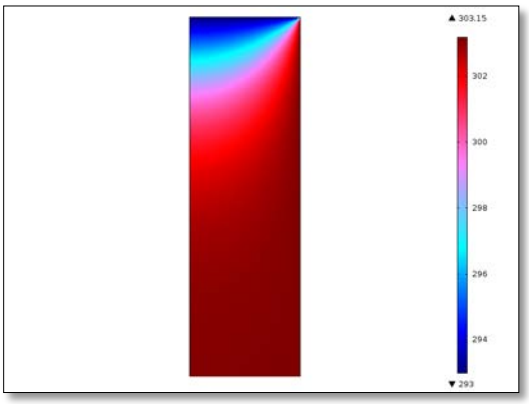

Figure 5. Steady state temperature profile of column matrix. A temperature gradient is only observed at the very beginning of the column, where temperature of the porous media equilibrates with the inflowing water. The column is kept at a constant temperature of 30℃ throughout the rest of its length.

The two pharmaceuticals involved in this study are creatine and creatinine, with retention factors 1.12 and 2.44 (Faulkner, 2013). The entrance of creatine in the first 5 seconds and the first 6mm of the column is shown in Figure 6. The inflow of species was specified to be 0.572 mol/m^3 for the very first second, after which a constant number of molecules propagated down the column. After 5 seconds, the species has started to diffuse and disperse in the porous matrix.

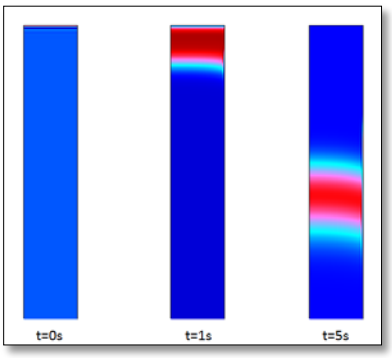

Figure 6. Concentration profile of creatine in the first 5 seconds. The concentration of creatine in the column is shown for t=0, 1 and 5 seconds, near the inlet of the column (6mm). Inflow is specified as 0.572 mol/m^3 for the 1st second. (Dark red represents where the relative highest concentration is; dark blue represents 0 mol/m^3.)

Figure 7 shows the transit of creatine and creatinine through the column. As shown in the figure, at t=10s, the solute molecules are closely clustered together; but as time goes on, dispersion increases, and the spread of their distribution inside the column increases. As the molecules move toward the outlet, the height of the concentration profile decreases while the width increases.

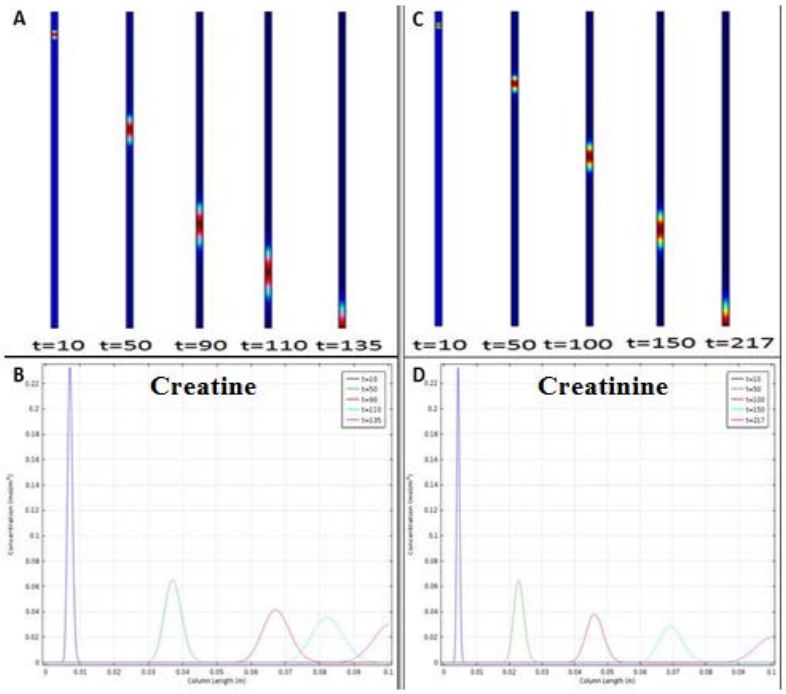

Figure 7. Concentration profiles of creatine and creatinine at selected time points up to the elution time. [A] and [C] show concentration surface plots of the column; [B] and [D] show the respective concentration profiles over the length of the column.

Given the goal of building model that simulates the separation and elution process of HPLC, a successful model must return a chromatogram-like plot that gives the separation of two molecular species. Figure 8 shows the resulting elution profiles of creatine and creatinine from the model developed in this study. The y-axis represents the average concentration of the species eluting from the HPLC column outlet over time.

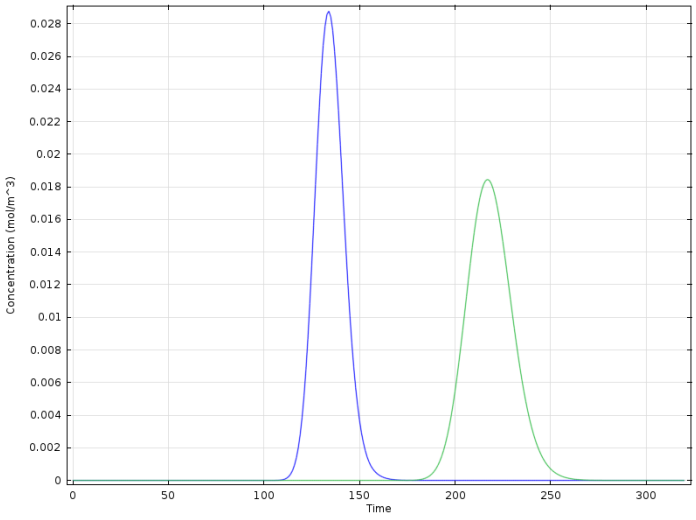

Figure 8. Elution profiles of creatine and creatinine using the mathematical model developed in this study.

It is clear from Figure 7 and Figure 8 that the column matrix retains creatinine longer than creatine, resulting in the later elution of creatinine. Creatine elutes at 134s (2.2 min) and creatinine elutes at 217 (4 min). This was expected since creatinine has a higher retention factor than creatine. The two peaks do not overlap, meaning that the two species are well separated on the chosen column. In order to measure the peak width, the plots were exported and analyzed using the programming software MATLAB. The peak widths of resulting elution profiles were measured at the half height of the peak (Full Width Half Maximum, FWHM), yielding widths of 17.3 seconds for creatine and 27.1 seconds for creatinine.

Verification & Accuracy Check

The results from the COMSOL model were compared to data provided by the column manufacturer, Thermo-Fisher (Faulkner, 2013), to validate the model. The geometry of the column and the parameters for the model were based off of the data provided by Thermo-Fisher. Calculated values for elution time and full width at half maximum (FWHM) were compared to the experimental values for both species in Figure 9.

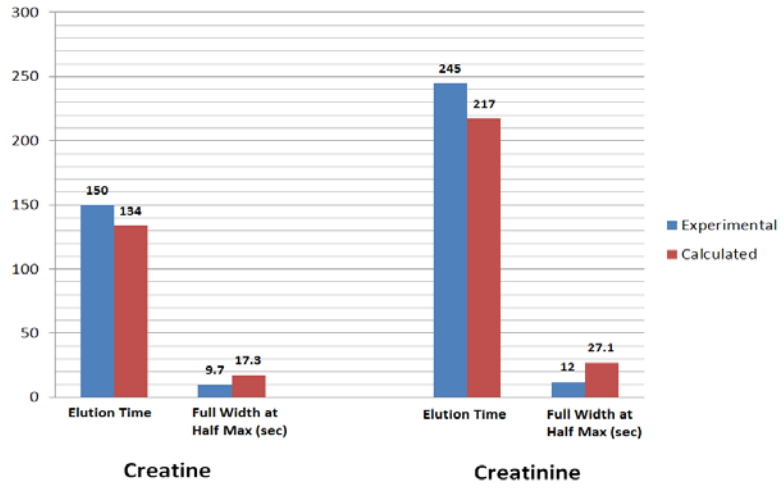

Figure 9. COMSOL HPLC model calculated data accuracy check. The model values for elution time and the elution peak's full width at half maximum for Creatine and Creatinine are compared side by side to the HPLC data provided by Thermo-Fisher.

As shown in the Figure 9, the calculated elution times of the model are reasonably close to the experimental elution times. A side by side comparison of the experimental and calculated elution profiles can be found in Appendix C, Figure A2. The calculated elution time for creatine was 10.67% lower than that of the experimental data. Similarly, calculated elution time for creatinine was 11.42% lower than that of the experimental data. The calculated FWHMs for the model were marginally but consistently greater than the experimental FWHMs. Despite the increased FWHM, the model still had the resolution needed to effectively separate the two species.

The discrepancies between experimental and calculated values can be attributed to the assumptions and approximations within the model. For example, the diffusivity used for the model was an effective diffusivity that combined the effects of molecular diffusion and dispersion into one value. The model also ignored the adsorption of the diluted species into the stationary phase. Instead, a single retention factor was implemented to approximate this physical process. Additionally, further complications in geometry are present in an actual HPLC set up, so an initial uniform concentration distribution at the top of the column may not be feasible. Given that the deviations of the model from experimental results are reasonably small, it is reasonable to conclude that the results from the model are valid.

Sensitivity Analysis

Sensitivity analysis was conducted to determine how results are impacted by the uncertainty of input parameters, with respect to the following parameters: column porosity, particle size, retention factor, and diffusivity. Column porosity was varied to reflect the common range of values for column porosities. The retention factor was varied by 10% to reflect the uncertainty provided by the column manufacturer [3]. Since the column's effective diffusivity was calculated from a mathematical approximation provided by Waters, a column manufacturer, this value was also varied by 20% [3]. Quantitative values were generated using the elution time and FWHM values for creatine.

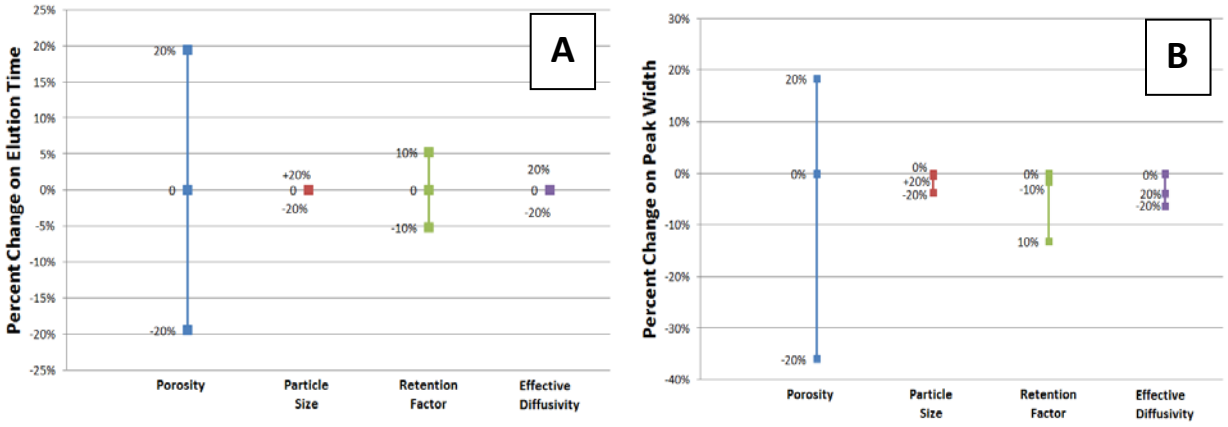

Figure 10. [A] Effect of varying parameters on creatine elution times. [B] Effect of varying parameters on creatine elution peak's FWHM. Percent variations displayed to the left of the data points.

Figure 10 reveals that the greatest sensitivity to change in the elution time and peak width is dependent on porosity and retention factor. The retention factor appears to vary linearly with elution time—which is plausible because it is a constant coefficient in the convection term of the species diffusion equation. Regarding the peak width, an increase in retention factor means that the solute spends more time in the column, which allows for less time to diffuse. According to Figure 10, an increase in retention factor results in a lower peak width. However, from a longer

elution time, one would expect the peak width to be increased. This discrepancy suggests a minor flaw in the model formulation - the retention factor is added as a coefficient in the velocity field model input under the Species Transport physics. However, the velocity field solution itself remains unchanged. In the variable definitions, the effective diffusivity is defined in terms of velocity only, when it should instead also be multiplied by the 1/(1+k') factor.

Regarding the effect of particle size on the peak width, a lower value would mean a more densely packed column, which produces more obstructions for the diffusing particles to traverse, which could explain the narrower peak. In analyzing small relative changes, an additional caveat lies in the FWHM Matlab program, which calculates the peak width based on discretized values, which is at best only an approximation.

The effect of changes in porosity and effective diffusivity are more interesting, as they are somewhat counterintuitive. The effective diffusivity has no influence on the elution time of the center, as expected, but both increasing and decreasing this value decrease the peak width. This change is small, so it is highly possible that there is a discrepancy in the baseline FWHM calculation due to approximation error. The other points do indicate a lower peak width corresponds to a lower diffusivity. Regarding the porosity, the more porous column generates a longer elution time. This is counterintuitive, as the solute particles should flow more easily through a highly porous matrix. The explanation, however, appears to lie in the computer model: in reality, the porosity and particle diameter of the column should be coupled, but in our model, both are specified as independent global parameters. The porosity only ever appears as a coefficient to the transient concentration term, whereas the particle size determines the diffusion of the solutes. This means that when porosity is changed, all other factors will remain the same, except for $\delta c/\delta t$, which will be transformed by $1/\varepsilon$. Thus a higher porosity results in a lower $\delta c/\delta t$, and a longer elution time. This point has exposed a potential weakness in our model: the porosity and column diameter are only linked through the literature. In future models, one value should be a function of the other, or vice versa.

Optimization

Several column parameters were varied to determine their effects on both the elution time of the solutes and the extent to which they have diffused inside the column. Peak separation refers to the difference in time (seconds) between when the two peaks of creatine and creatinine concentration occur. In general a high value of peak separation is desired, signifying that the creatine and creatinine are very effectively separated. Low FWHM values are desired, as this signifies low dispersion of the solutes inside the column. If dispersion values are too high, then the creatine and creatinine plots are more likely to intersect on the plots and remain mixed, which indicates a failure to separate the species.

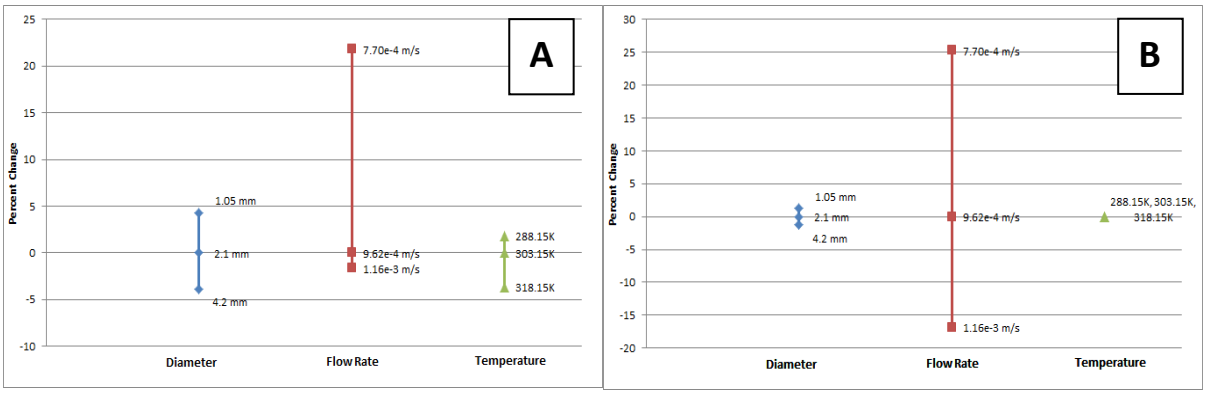

Figure 11. [A] Effect of varying parameters on peak separation. [B] Effect of varying parameters on average FWHM. Parameter values to the left of the data points.

According to the Figure 11, flow rates have a large effect on the separation of creatine and creatinine, but do not influence the dispersion of the solutes. However, there is a small upward trend in FWHM at lower flow velocities because the solutes spend more time inside the column and have more time to diffuse axially. The peak separation does vary widely across the tested values, decreasing rapidly as the flow velocity is increased. This is expected because if the solutes are being pushed through the column with great force, the solutes have less time to interact with the solid phase, which is the source of their peak separation. For a full optimization, a flow rate must be specified that is low enough to guarantee high peak separation, but just fast enough so that a large increase in solute dispersion is not seen. Of the tested values, it appears that 7.70e-4 m/s performs better than the manufacturer-suggested flow velocity of 9.624e-4.

Increasing the length of the column will tend to increase the mean separation because the solutes spend more time differentially interacting with the solid phase. However, the more time solutes spend inside the column, the more they diffuse and disperse, so that at some point the elution profiles may mix together again despite large mean separations. Based on the Figure 10, it appears that at higher lengths in this range, the peak separation improves dramatically with a small, but consistent, rise in solute dispersion, so the column length can be increased by a large margin before the tradeoff point occurs.

The column diameter appears to have little effect on the peak separation, though this may be unique to this model because the concentration at the inflow is not fixed, but uniform across the radius. This can be altered in future models to test differences in diameter. Higher dispersions at higher temperatures may be due to the solute flow retardation at the boundary. With a wider column, there is more potential for a non-uniform solute concentration profile to form along the radial direction.

As expected, varying the heating of the column outer boundary temperature did not have much effect on either peak separation or solute dispersion. As the model is designed currently,

lower temperatures should be favored, as this result in lower mass diffusivities and thus lower FWHM values.

The effect on peak separation by temperature in the optimization model does not logically make sense at first glance: diffusivity is linearly correlated with temperature, so higher column temperatures should produce greater diffusivities, and thus greater separation. This however, is exactly the opposite of what results in Figure 11A. Figure 12 shows that this phenomenon is indeed due to the effective diffusivity: the lowest temperature produces the highest effective diffusivity. Conversely, the traditional diffusivities that correspond to these temperatures increase with temperature as expected: 4.6e-10, 7.0e-10, and 9.9e-10 m^2/s.

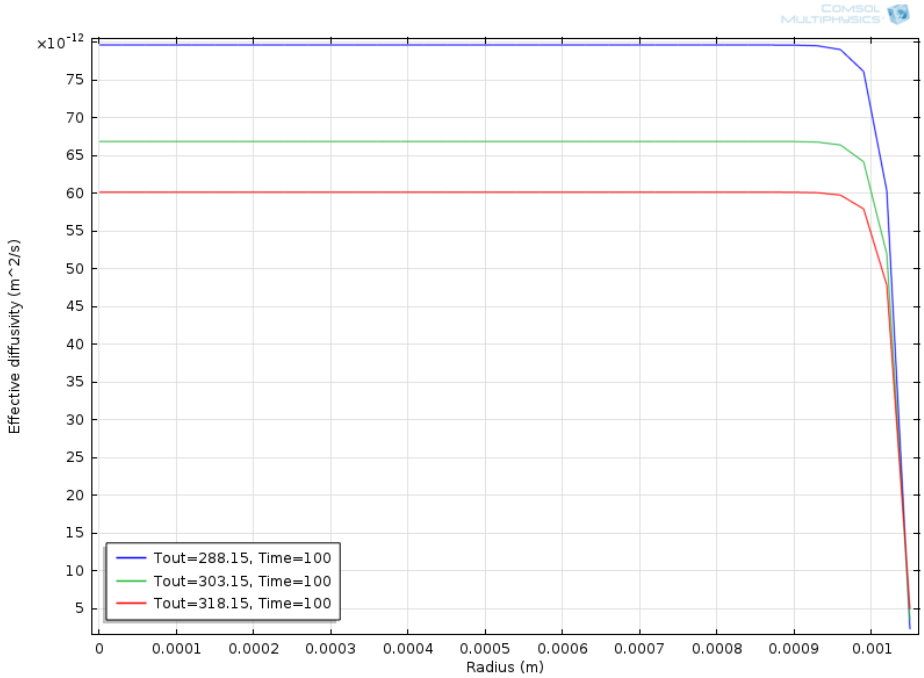

Figure 12. The effective diffusivity profile along the radius of the column, taken at the midpoint of the column after 100s at steady state. Lower temperatures tend to increase the effective diffusivity.

In the equation for D_{eff} presented in the schematic, D_{mol} appears in the denominator of the third velocity-dependent term. At a certain particle size and fluid velocity, this term will actually overpower the second term with D_{mol} in the numerator, making D_{eff} an inverse function of D_{mol}. This is shown in Appendix C, Figure C. Physically, this relationship occurs because the solute particles have a tendency to interact with the solid matrix particles, and lag behind the flow. This interaction is always at equilibrium, so only a fraction are slowed interacting with the solid, causing the solute particles as a whole to spread out more and resulting in a higher elution peak width. The third term of the effective diffusivity compensates for this effect by increasing the diffusivity itself. However, molecules with a high traditional diffusivity will be able to more freely diffuse around the mobile liquid phase, making them less available to interact with the solid. Thus high D_{mol} molecules are slowed less by the solid phase. In the particular column diameter and flow rate parameters selected for this column, D_{mol} and D_{eff} are actually inversely

related because this phenomenon dominates over the more traditional longitudinal diffusion term.

Optimization of the column using an objective function

For a more quantitative method of optimization, an objective function was used to find an optimal length for the column. Two measures of HPLC effectiveness are resolution and theoretical plate height (TPH). A description of these functions can be found in Appendix C, Table A4.

The resolution and theoretical plate heights for a range of column lengths were calculated for a range of column lengths from 50 mm to 150 mm. The elution times and peak widths were plotted against the range of column lengths to establish equations for these functions in terms of column lengths, as displayed in Figure 13.

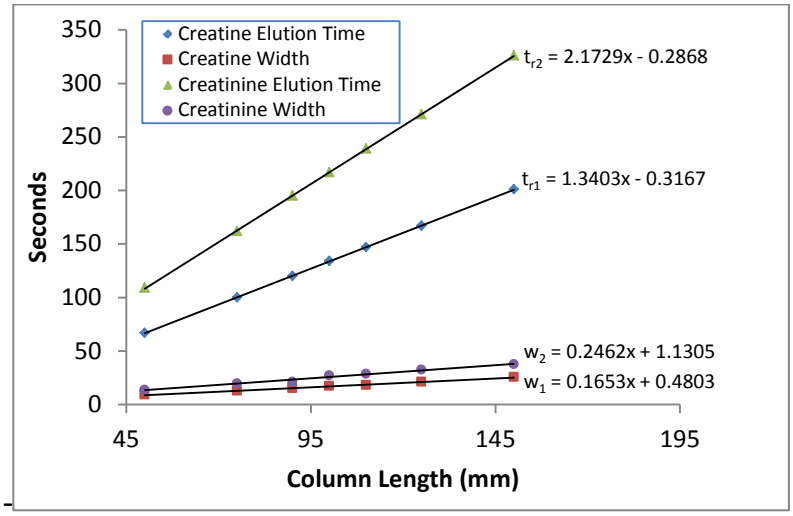

Figure 13. Functions for elution time and peak width in terms of column length. Using a range of column lengths from 50 to 150 mm, a linear fit was used to establish the elution time and peak width as a function of column length

After finding elution times and peak widths in terms of column length, these functions were used to find resolution and TPH. For the objective function, the resolution, TPH for creatine, and TPH for creatinine are standardized by average their values and summed. In addition, the increasing elution time was penalized.

Objective function: $J(x) = \frac{Resolution(x)}{3.88} + \frac{THP_{creatine}(x)}{981} + \frac{THP_{creatinine}(x)}{1131} - 0.15\left(\frac{t_{r1}}{134} - \frac{t_{r2}}{217}\right)$

This objective function will be maximized to find the optimal column length. The objective function J is plotted for the range of column lengths is seen in Figure 14.

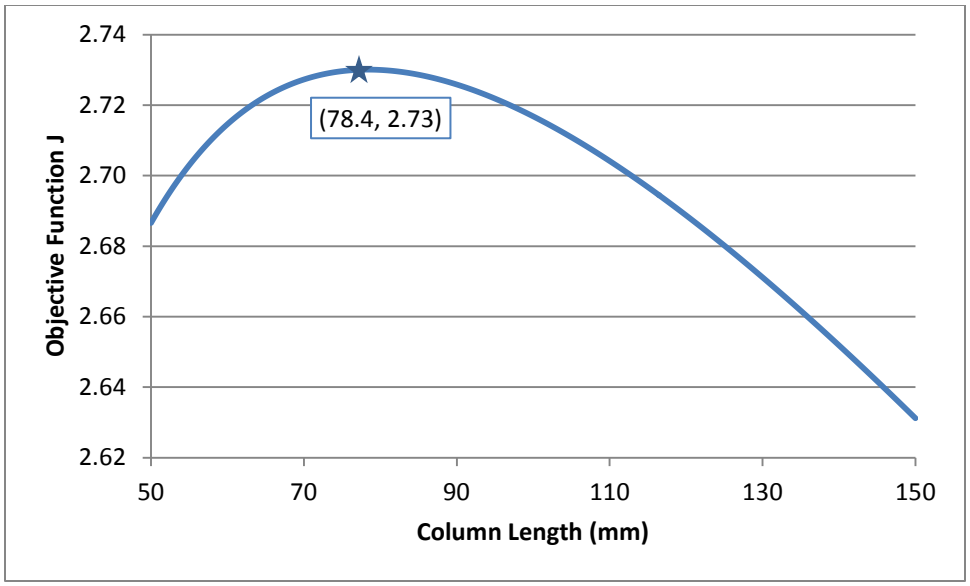

Figure 14. Objective function plotted for a range of column lengths. The maximum point, and therefore optimal column length, was 78.4 mm.

As seen in Figure 14, a column length of 78.4 mm was found to be the optimal column length. This process could be repeated to find optimal values for other parameters of the column.

Conclusion and Design Recommendations

In this study, a mathematical model was successfully developed in COMSOL Multiphysics to simulate a High Pressure Liquid Chromatography (HPLC) process for the separation of the compounds creatine and creatinine. The resulting elution profiles of creatine and creatinine obtained from this model exhibited sharp peaks and high separation from each other, and their elution times closely matched the chromatograms obtained by experiment. An optimization analysis conducted on various parameters as they relate to this column found that an optimal length of around 80 mm was ideal for this separation process, and the manufacturer-specified diameter and flow rate are already roughly optimal. Additionally, an interesting inverse relationship between effective and actual diffusion can develop under certain velocity and particle size conditions due to solid matrix interactions, as in this case. This interesting property of the Van Deemter model shows that the column temperature must be carefully considered in future models and experiments to accurately predict species diffusion behavior.

Although this model works very well for the elution of creatine and creatinine on the Hypercarb Reverse Phase HPLC column by Thermo Fisher Scientific, caution must be taken in applying it to other HPLC experiments. The elution process was simplified for this model, and many physical parameters are interrelated in ways that cannot be captured by simplified mathematical expressions. For example, the retention factor depends on the column matrix, experimental conditions and the compound itself, all of which is poorly characterized by one dimensionless constant. Additionally, the porosity and particle size of the solid matrix are interrelated, and changing one without the other causes unexpected results as seen in the sensitivity analysis.

The situational nature of many of these parameters means that new models must be developed to accurately model new elution processes. The development cost of this endeavor may serve as an economic constraint on pharmaceutical companies, especially when it is traditional to depend only on physical tests. Additionally, matrix properties such as porosity, permeability, and density were represented as simple numerical values in this model, and the manufacturing process may not consistently produce columns which match these idealized figures.

Despite these constraints, the mathematical model developed in this study can still provide design recommendations and serve as a useful tool in analytical chemistry before scientists purchase columns to perform an HPLC experiment. Using computer modeling, the tradeoff between desired and undesired effects can be carefully controlled in a quantitative manner, and thus maximize the physical performance. Even if this particular separation process is never conducted, this work serves as a general guide for designing other solute-specific designs and demonstrates that the resulting models are able to closely reproduce experimental data. As a single HPLC column is often priced between $200 and $1000 (Ryan, 2013), testing series of columns without the guidance of computer simulations can be cost-prohibitive, as well as labor intensive and time consuming. Mathematical models such as the one described in this work will help pharmaceutical companies to perform product testing at a much higher efficiency and lower cost.

Appendix

Appendix A. Mathematical Statement

1. Steady-state fluid flow
    a. The fluid flow is described by the Navier-Stokes equation:

$$\rho(\frac{\delta v}{\delta t} + v\nabla v) = -\nabla p + \mu\nabla^2 v$$

In the case of flow through a very small space, the viscous forces far outweigh the inertial forces, and the change in velocity is expected to be small, so the terms on the left side of the above equation are assumed to be negligible. This steady state situation is typically termed Stokes flow. Additionally, Darcy's law is used to describe the flow through the solid matrix:

$$v = -\frac{k}{\mu}\nabla p$$

where the constant k is the permeability, determined by the solid matrix particle size and the porosity:

$$k = \frac{d_p^2 \varepsilon^3}{180(1-\varepsilon)^2}$$

Combining the Darcy expression with the assumptions describing Navier-Stokes flow above results in the Brinkman equation, to describe the fluid flow through a porous medium:

$$\mu\nabla^2 v - \frac{\mu}{k}v - \nabla p = 0$$

This equation is run using cylindrical coordinates, and pressure is applied at the inlet only in the axial direction, resulting in the steady state

$$\mu\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial v_z}{\partial r}\right) + \frac{\partial^2 v_z}{\partial z^2}\right] - \frac{\mu}{k}v_z - \frac{\partial P}{\partial z} = 0$$

2. Steady-state column heating
    a. The inflowing fluid is at a constant temperature at the boundary of 20 °C, and the outside wall of the column is kept at a constant temperature of 30°C. This situation is described by the traditional steady state heat equation:

$$\rho C_p v_z \frac{\partial T}{\partial z} = k\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial T}{\partial r}\right) + \frac{\partial^2 T}{\partial z^2}\right]$$

The temperature of the fluid influences the diffusivity expression, which is described below in (3), as well as four fluid properties, density, viscosity, heat capacity, and thermal conductivity, all of which are described by polynomial functions listed in the Appendix A, Table A3.

3. Creatine and creatinine diffusion and dispersion in the column
    a. The diffusion of the two solutes is described using the transient species transport equation:

$$\frac{\partial c}{\partial t} + \nabla(uc) = \nabla(D\nabla c)$$

To account for interactions with the solid matrix and the dispersion caused by the flow, the following terms are added to the equation in cylindrical coordinates (COMSOL Model Documentation, 2012):

$$(\varepsilon)\frac{\partial c}{\partial t} + \frac{1}{1+k'}v_z\frac{\partial c}{\partial z} = D_{eff}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial c}{\partial r}\right) + \frac{\partial^2 c}{\partial z^2}\right]$$

Here ε is the porosity of the solid matrix, and the retention factor k' is a constant describing how much the solutes are slowed by the matrix relative to the fluid flow. The retention factors for creatine and creatinine were obtained from a specification sheet of the column from the manufacturer, Thermo-Fisher. The effective diffusion coefficient modifies the diffusivity (D) by taking into account the dispersion using the van Deemter equation (Bunner, 2008), which is dependent on the velocity:

$$\mathrm{D_{eff}} = \frac{\mathrm{u}}{2}\left(1.5d_p + \frac{D}{u} + \frac{d_p^2}{6D}u\right)$$

Appendix B. Property Tables

Table A1. HPLC Column Properties.

Parameter	Values	Source
Column Diameter	2.1 mm	(Faulnker, 2012)
Packing Particle Size (d_p)	5 µm	(Faulnker, 2012)
Column Length	100 mm	(Faulnker, 2012)

Table A2. Comsol Parameters.

Parameter	Value	Source
Creatine Retention Factor (k')	1.12	(Faulker, 2012)
Creatinine Retention Factor	2.44	(Faulker, 2012)
Porosity (ε)	0.6	(COMSOL Model Documentation, 2012)

Flow rate	9.624E-4 m/s	(Faulker, 2012)
Column Permeability	1.875E-13 m^2	(Bunner, 2008)
Molecular Weight of Creatine	131.13 g/mol	(PubChem)
Molecular Weight of Creatinine	113.1179 g/mol	(PubChem)
Density of Packing Material (ρ)	2200.0 kg/m^3	(Graphite)
Viscosity of Water (μ)	8.9E-4 Pa·s	(EngineeringToolbox)
Thermal Conductivity of Porous Matrix	10 W/(m*K)	(Jobmanna, 2009)
Specific Heat Capacity of Porous Matrix	710 J/(kg*K)	(EngineeringToolbox)

Table A3. *COMSOL Variables.*

Variable	Equation	Source
Permeability	$K = \frac{{d_p}^2 \times \varepsilon^3}{180(1-\varepsilon)^2}$	(Bunner, 2008)
Molecular Diffusivity	$D_{A,water} = \frac{9.40 \times 10^{-15} * T}{\mu {M_w}^{1/3}}$	(Saltzman)
Plate Height	$H(u) = 1.5d_p + \frac{D_{mol}}{u} + \frac{{d_p}^2}{6D_{mol}}u$	(Bunner, 2008)
Effective Diffusivity	$D_{eff} = \frac{H(u) * u}{2}$	(Bunner, 2008)
Bulk Density of Packed Bed	$\rho_b = \rho(1-\varepsilon)$	(COMSOL Model Documentation, 2012)
Temperature Dependent Viscosity of Water	$\mu(T) = 0.0967 - 8.207 \times 10^{-4} * T + 2.344 \times 10^{-6} * T^2 - 2.244 \times 10^{-9} * T^3$	(Pramuditya, 2011)
Temperature Dependent Density of Water	$\rho(T) = 765.33 + 1.8142T - 0.0035T^2$	(Pramuditya, 2011)
Temperature Dependent Heat Capacity of Water	$C_p(T) = 28.07 - 0.2817T + 1.25 \times 10^{-3}T^2 - 2.48 \times 10^{-6}T^3 + 1.857 \times 10^{-9}T^4$	(Pramuditya, 2011)
Temperature Dependent Thermal Conductivity of Water	$k(T) = -0.5752 + 6.397 \times 10^{-3}T - 8.151 \times 10^{-6}T^2$	(Pramuditya, 2011)
Volume Fraction	$\theta_P = 1 - \varepsilon$	(Henden, 2006)

Appendix C. Solution Strategy

COMSOL Solver

Direct solver was used in this study to solve for the steady state fluid flow, steady state heat transfer and time dependent species transfer through the porous matrix. For the time dependent species transfer module, a step size of 1s was used. Steps taken by solver were set to be strict. A relative tolerance of 0.01 and an absolute tolerance of 1×10^{-3} were used by the time-stepping algorithm in the COMSOL. Since the absolute tolerance is much smaller than the relative tolerance, the relative tolerance will be the ruling criterion for error control. The time-dependent solver will converge and move onto the next time step when the following inequality is satisfied (Support Knowledge Base, 2013):

$$|\epsilon_i| \leq \text{Tolerance}_{\text{relative}}*|v_i| + \text{Tolerance}_{\text{Absolute}*}|v_i|$$

where the ϵ_i is the error for the i^{th} time step, and v_i is the value of the dependent variable for the i^{th} time step. For the variable of concentration, c, this inequality would be:

$$|c_i - c_{i-1}| \leq \text{Tolerance}_{\text{relative}}*|c_i| + \text{Tolerance}_{\text{Absolute}*}|c_i|$$

Mesh Convergence

Structured mesh was used in this study, with 35 elements along the radial direction and 2200 elements along the axial direction (77000 elements in total). Due to the length (100mm) of the column being much longer than the radius (1.05mm), only a small portion of the column with mesh is shown in Figure A1 [A]. Mesh convergence was determined by plotting the outlet concentration after 150s (2.5min) over the number of elements. This specific time point was selected since it was the manufacturer's given elution time for creatine. At this time point, there would be a significant concentration of creatine flowing through the outlet. Performing a mesh convergence would result in an appropriate mesh with minimal discretization error. As seen in Figure A1 [B], the concentration stops changing after 63,000 elements, at which point the solution converges. This proved that that choice of 77,000 elements was appropriate in the solution process.

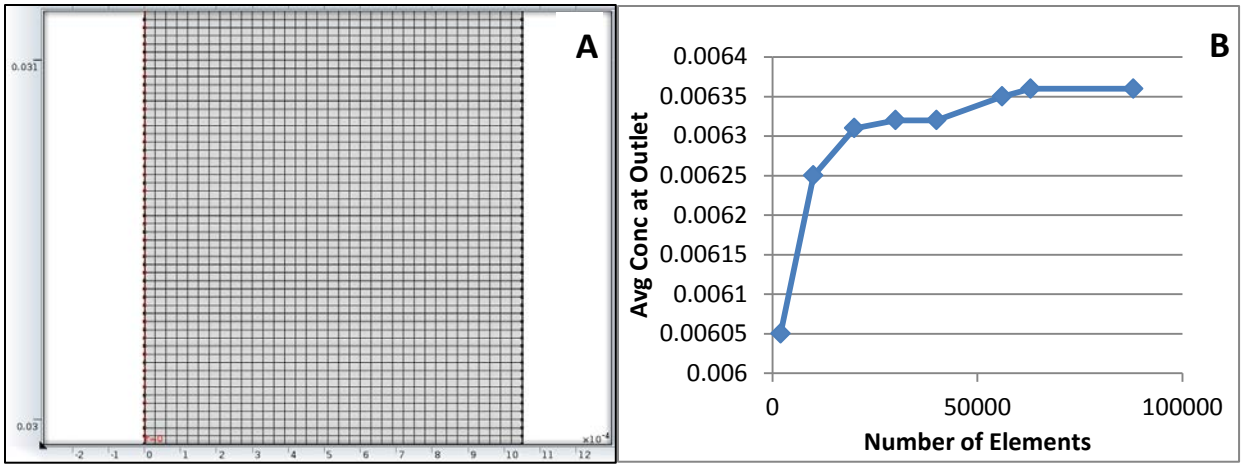

Figure A1. Element mesh [A] and mesh convergence plot [B]. [A] Structured mesh was used in this study to build the column model. [B] Mesh convergence was evaluated by the average creatine concentration at the outlet at 150s; it converged at 63,000 elements.

Appendix D. Additional Visuals

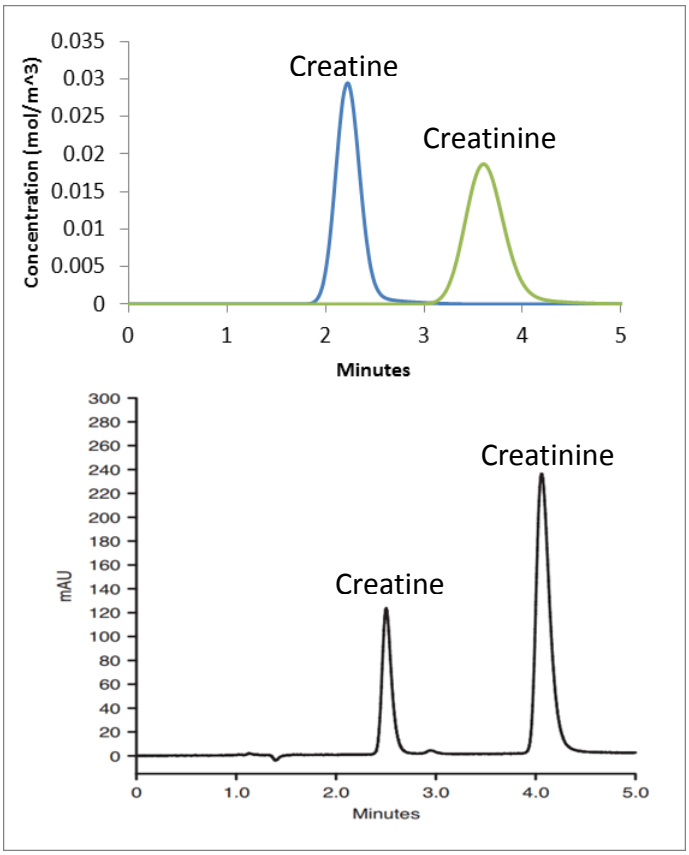

Figure A2. Comparison between elution profiles from this study and actual experiment.

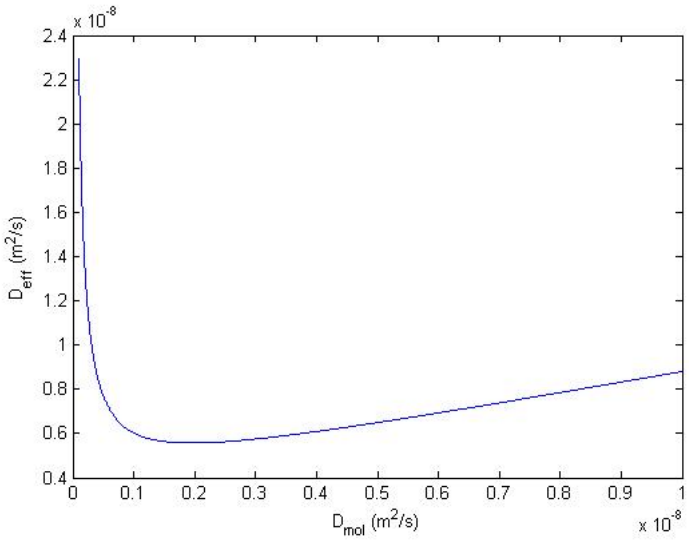

Figure A3. The relationship between effective and traditional diffusivities according to Van Deemter's equation, given a solid phase particle diameter of $5*10^{-6}$ m and a flow rate of $9.6*10^{-4}$ m/s. Below a certain threshold, D_{eff} actually varies inversely with D_{mol} due to additional band broadening caused by increased interactions with the solid matrix, which is determined by the solid matrix particle size, and the flow rate.

Table A4. A description of resolution and Theoretical Plate Height

Function	Equation	Description
Resolution	$R = \frac{(t_{r2} - t_{r1})}{1/2(w_1 + w_2)}$ where t_{r1} and t_{r2} and w_1 and w_2 are the times and widths, respectively, of the two adjacent peaks	A measure of separation between the two species being separated
Theoretical Plate Height	$N = 16\left(\frac{t_r}{w}\right)^2$ where t_r is the elution time and w is the peak width	A measure of HPLC effectiveness, peak clarity

Appendix E.

References

New Drug Approval Process. (2013). *Drug Information Online*. Accessed on March 10, 2013. http://www.drugs.com/fda-approval-process.html

Waters.(2013). *HPLC-High Pressure Liquid Chromatography*. Milford, MA. Accessed on March 10, 2013. http://www.waters.com/waters/nav.htm?locale=en_US&cid=10048919

Faulkner, W. (2012). Analysis of Creatine and Creatinine on a Porous Graphitic Carbon Column by HPLC/UV. Thermo Fisher Scientific. Application Note 20512.

MacNeil, L. et al. (2005). Analysis of creatine, creatinine, creatine-d3 and creatinine-d3 in urine, plasma, and red blood cells by HPLC and GC-MS to follow the fate of ingested creatine-d3. *Journal of Chromatography B*. 827:210-215.

NSN Publishing. (2013). Creatine Information Center. Accessed on May 2, 2013. http://www.creatinemonohydrate.net/illustrations-reactions

Cobert Associates. (2013). How to Choose an HPLC Column. Accessed on May 2, 2013. http://cobertassociates.com/How%20to%20Choose%20an%20HPLC%20Column.htm

COMSOL Model Documentation. (2012). Liquid Chromatography.

Bunner, B. et al. (2008). Simulation of Chromatographic Band Transport. Excerpt from the Proceedings of the COMSOL Conference 2008 Boston.

Gottlieb, S. and Hosfelt, J. Chromatography. UC Davis Chem Wiki. Accessed on May 2, 2013. http://chemwiki.ucdavis.edu/Analytical_Chemistry/Instrumental_Analysis/Chromatography

Kazakhevich, Y. (2010). HPLC for Pharmaceutical Scientists. Accessed on May 2, 2013. http://hplc.chem.shu.edu/NEW/HPLC_Book/Theory/th_vandm.html

Ryan, David. (2013). *Chapter 28: High-Performance Liquid Chromatography (HPLC)*. University of Massachusetts.Lowell, MA. Retrieved Apr 6 2013. http://faculty.uml.edu/david_ryan/84.314/Instrumental%20Lecture%2018.pdf

PubChem (n.d.). Creatinine - PubChem. (n.d.). Accessed on May 2, 2013. http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=588

PubChem (n.d.). Creatine - PubChem. Accessed on May 2, 2013. http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=586

Graphite Mineral Data. (n.d.). Accessed on May 2, 2013. http://webmineral.com/data/Graphite.shtml

EngineeringToolbox (n.d.). Water - Dynamic and Kinematic Viscosity. Accessed on May 2, 2013. http://www.engineeringtoolbox.com/water-dynamic-kinematic-viscosity-d_596.html

Pramuditya, S. (2011). Water Thermodynamic Properties « Hey, what's going on?. Accessed on May 2, 2013. http://syeilendrapramuditya.wordpress.com/2011/08/20/water-thermodynamic-properties/

Henden. (2006). Porosity and volume fraction. Accessed on May 2, 2013. http://www.pvv.org/~perchrh/fibforsk/thesis/node19.html

Jobmanna , M. (2009). Influence of graphite and quartz addition on the thermo–physical properties of bentonite for sealing heat-generating radioactive waste. Applied Clay Science, (44), 206-210.

EngineeringToolbox (n.d.). Solids - Specific Heats. Accessed on May 2, 2013. http://www.engineeringtoolbox.com/specific-heat-solids-d_154.html

Smith-Palmer, T. (2002). Separation methods applicable to urinary creatine and creatinine. *Journal of Chromatography B*. 781(1-2): 93-106.

Support Knowledge Base. Convergence criterion for a time dependent problem. Multiphysics Modeling and Simulation Software. Accessed May 2, 2013. http://www.comsol.com/community/forums/general/thread/13859/